Acquired genetic and epigenetic variation in human pluripotent stem cells

Kyriakides O, Halliwell J.A., Andrews P.W.

Contents

Abbreviations

aCGH- array comparative genome hybridisation

CNV- copy number variation

FISH- fluorescent in situ hybridisation

hESC- human embryonic stem cell

hiPSC- human induced pluripotent stem cell

hPSC- human pluripotent stem cell

NGS- next generation sequencing

qPCR- quantitative polymerase chain reaction

SNP- single nucleotide polymorphism

TGCT- testicular germ cell tumour

Abstract

Human pluripotent stem cells (hPSCs) can acquire non-random genomic variation during culture. Some of these changes are common in tumours and confer a selective growth advantage in culture. Additionally there is evidence that reprogramming of human induced pluripotent stem cells (hiPSCs) introduces mutations as well. This poses a challenge to both the safety of clinical applications, and the reliability of basic research using hPSCs carrying genomic variation. A number of methods are available for monitoring the genomic integrity of hPSCs and a balance between practicality and sensitivity must be considered to choose the appropriate methods for each use of hPSCs. Adjusting protocols by which hPSCs are derived and cultured is an evolving process that is important to minimise acquired genomic variation. Finally, assessing genetic variation for its potential impact is becoming increasingly important as techniques to detect genome wide variation improve.

Introduction

Human pluripotent stem cells (hPSCs) can be derived from embryos or induced from somatic cells (Takahashi et al., 2007, Yu et al., 2007, Thomson et al., 1998). These cells have the ability to produce cell types from any of the three germ layers and can self renew. Excitement surrounding hPSCs is fuelled by potential uses in studying development, modelling disease and regenerative medicine.

Taking Parkinson's disease as an example, disease models have been developed by reprogramming patients' fibroblasts to human induced pluripotent stem cells (hiPSCs) facilitating a better understanding of the Parkinson's disease genotype (Soldner et al., 2009). Furthermore, through developing protocols for the differentiation of hPSCs to dopaminergic neurons a greater understanding of neuronal development cues has gradually built up (Perrier et al., 2004). This knowledge then allows the

gradual translation into regenerative medicine treatments (Kriks et al., 2011).

Similarly using tissue from long QT patients, hiPSC derived cardiac myocytes have been generated which show a characteristic reduction in the delayed rectifier potassium current (Itzhaki et al., 2011). Furthermore the long QT hiPSC derived cardiac myocyte model was used to screen for pharmacological agents providing an improvement to the phenotype (Itzhaki et al., 2011).

These examples demonstrate the importance of hPSC research in a wide array of fields. However, in order to realise this potential, hPSCs must be maintained, often in large numbers, in culture. hPSCs show apparent immortal self-renewal in culture which distinguishes them from their in vivo embryonic counterparts which quickly become restricted in fate (Andrews, 2002). Since their first derivation, numerous studies have shown that hPSCs are subject to genomic change within culture. Furthermore, hiPSCs show additional signs of genetic instability associated with the reprogramming process.

This review aims to first summarise the current knowledge on acquired genomic change in hPSCs. Following this are discussions on emerging approaches to monitor, minimise and assess genomic change that are important considerations in the field of hPSC research.

Genetic change in human pluripotent stem cell culture

Genetic change can occur spontaneously in any cell but, through a combination of natural senescence and apoptosis, most never become established within the overall population. Indeed, post mortem neural tissue shows low level mosaicism which may help to produce functional diversity (Rehen et al., 2005) and normal pluripotent stem cell populations are likewise chromosomally heterogeneous (Peterson et al., 2011). However, some hPSC cultures show non-random genetic change that can come

to dominate the population (Draper et al., 2004). Commonly these involve gains of parts of chromosomes 1, 12, 17 and 20 and losses of regions of chromosomes 10, 18 and 22 (Figure 1) (Amps et al., 2011).

For example, in one study over 50% of 30 human embryonic stem cell (hESC) lines maintained over 18 months developed karyotype abnormalities, with 17q and chromosome 12 trisomy being the most frequent change (Baker et al., 2007). Furthermore, the same karyotype abnormalities have been reported independently in other lines (Cowan et al., 2004, Mitalipova et al., 2005). In all cases the abnormalities were observed only after continued culture.

These changes are not exclusive to hESCs. Using a technique that infers karyotype abnormalities from gene expression data, it was shown that genes on chromosome 12 were consistently overexpressed in hiPSC lines as well (Mayshar et al., 2010). Together these data imply that the genetic aberrations observed are characteristic of pluripotent stem cell culture, rather than the source of the cells from either embryo or fibroblast.

In a large scale screening of 125 hESC lines by single-nucleotide polymorphism (SNP) array analysis a sub-chromosomal copy number gain of part of the long arm of chromosome 20 (20q11.21) was identified in twenty-two cell lines (Amps et al., 2011). In all cases the duplications overlapped, sharing a minimal amplicon region of 0.55MB pairs, and the same copy number variant (CNV) has been identified independently in both hESCs and hiPSCs (Laurent et al., 2011).

Three genes within the 20q11.21 minimal amplicon are commonly expressed in hESCs. One of which, BCL2L1, forms two alternative transcripts that encode both a pro-apoptotic protein and an anti-apoptotic protein. In embryonic stem cells it is almost exclusively the anti-apoptotic protein, BCL-XL, that is expressed (Avery et al., 2013).

This finding gives support to the hypothesis that the non-random genetic change observed within hPSCs is driven by selection resulting in advantageous genetic variation becoming widespread over long term culture. Using the 20q11.21 CNV as an

example, we can assume this arises randomly and because of the unlimited proliferative potential of hPSCs the extra dose of BCL-XL conferred by this CNV could confer a selective advantage through its anti-apoptotic effects. Therefore, continuous passaging of a cell culture carrying this CNV will lead to its gradual accumulation within the cell population.

Experimental evidence for this model has been provided through the comparison of hESC lines carrying the 20q11.21 CNV with control hESC lines (Avery et al., 2013). In this study, population doubling times of 35 hours and 138 hours were reported respectively. Flow cytometry showed no difference in the distribution of cells throughout the cell cycle within each population and time-lapse confocal microscopy confirmed a similar absolute cell division time. This indicates that the reduced population doubling time observed in 20q11.21 CNV carrying cells was due to a reduction in apoptosis rather than an increase in proliferation. The action of BCL-XL specifically in this process was confirmed by overexpressing only BCL-XL in a separate cell line which mirrored the results of the cells carrying the whole 20q11.21 CNV.

Strikingly, this process by which cells can acquire a growth advantage upon prolonged culture closely resembles aspects of tumorigenesis (Figure 2), which is likewise thought to originate from mutations in a single cell that allow it to escape from tight growth control leading to selective clonal expansion (Fialkow et al., 1977). It is therefore possible that culture adaptation is an in vitro mimicry of this micro-evolutionary process. This raises concerns for the clinical application of hPSCs as it is plausible for such genetic change to confer malignant properties. For example isochromosome of 12p is used as a clinical marker of testicular germ cell tumours (TGCT) (Bosl et al., 1989). Furthermore, fluorescent in situ hybridisation (FISH) in human embryonal carcinoma cells, the malignant counterpart to hESCs, found that 6/9 carried the 20q11.21 amplification (Avery et al., 2013), which may suggest it can similarly drive growth advantage in malignant cells.

Genetic change has been detected on every chromosome, although aberrations on chromosome 4 are exceptionally rare, during hPSC culture (Amps et al., 2011). However, a so far unanswered question is why particular aberrations, such as those on chromosomes 12, 17 and 20, are so common? Recently it has been demonstrated that under replicative stress hESCs fail to activate key proteins, such as CHK1 and ATR, involved in the S-phase checkpoint despite normal levels of expression (Desmarais et al., 2012). Furthermore hESCs showed an upregulation in apoptotic markers and caspase 3 activation. This suggests that an intrinsic characteristic of hESCs is to eliminate cells acquiring DNA damage without an attempt at repair. This may be a desirable mechanism for protecting genome integrity as genetic change in ESCs in vivo would be passed on to the whole organism and could prove catastrophic. These findings have relevance to the discussion of acquired genomic variation. If hESCs normally protect genomic integrity through apoptosis rather than DNA repair then an acquired variation such as the 20q11.21 CNV would provide a particular selective advantage. The resistance to apoptosis conferred by the extra dose of BCL-XL in cells carrying this CNV is likely to help them thrive under these conditions. This could go some way to explaining why chromosome 20q variations develop so commonly in hPSC culture.

Thus far similar evidence for driving genes on chromosomes 12 and 17 has been elusive. This is largely due to the scale of the changes. The 20q minimal amplicon is only 0.55MB, so presented a limited number of candidate genes to investigate. In contrast the changes in chromosomes 12 and 17 usually involve a duplication of either the whole chromosome or an arm so pinpointing the driving genes involved is more difficult.

Nevertheless, candidate genes have been suggested. For example the gene BIRC5, located at 17q25.3, is known to have anti-apoptotic properties and is highly expressed in teratomas, the tumours formed by hESCs (Blum et al., 2009). Likewise NANOG, found at 12p13.31, contributes to maintaining pluripotency (Mitsui et al.,

2003), so if overexpressed may make cells more likely to continue self-renewal. However, detailed analysis shows the closest minimal amplicon falls upstream of NANOG on its unexpressed pseudogene (Laurent et al., 2011). Furthermore, the same minimal amplicon was found to be just as prevalent in the reference samples (Amps et al., 2011) and so is unlikely to be causing a change in cell behaviour.

However it is important not to dismiss the possibility that the phenotypic growth advantage conferred by these chromosomal aberrations is a result of a change in expression of multiple genes. This could explain why genetic change involving these chromosomes tends to involve whole or large duplications.

Epigenetic change

The epigenetic status of a cell is highly important in gene expression and therefore dictating its specific phenotype (Ferguson-Smith, 2011). Particularly relevant are the processes of genome imprinting whereby DNA methylation patterns produce monoallelic expression of particular genes in a parent of origin manner (Butler, 2009). Previous studies had observed epigenetic instability in cultured mouse ESC (Humpherys et al., 2001) and hypothesised a link between assisted reproductive technology and epigenetic disorders (Butler, 2009). This prompted investigations into whether removing hESCs from their in vivo environment and prolonged culture could perturb epigenetic imprinting.

In an early study of six imprinted genes in four hESC lines the normally paternally imprinted gene, H19, gained biallelic expression on prolonged culture (Rugg-Gunn et al., 2005). The H19 gene stands out from the other five genes investigated since it acquires methylation during embryonic development. However, upon closer inspection the re-expressed allele of H19 still showed methylation typical of an imprinted gene suggesting the re-expression must occur through an alternative mechanism (Rugg-Gunn et al., 2005).

In further studies of over 2000 loci by restriction landmark genome scanning all six hESC lines showed high levels of epigenetic instability which was reliably fixed within the cell population (Allegrucci et al., 2007). Another study found that IGF2 became biallelically expressed in an hESC line grown by one laboratory, whereas cultures of the same line grown by a different laboratory did not exhibit the same biallelic expression, which may suggest that culture conditions can have an effect on the epigenetic status of cultured hESCs (Rugg-Gunn et al., 2007).

A more recent study of 205 hPSCs and 130 somatic samples provided interesting insight into tissue specific versus pluripotent epigenetic character (Nazor et al., 2012). This study also detailed the correlation between either hypermethylation or hypomethylation with the loss of allele specific expression of numerous genes in hPSC's. Additionally the group reported that in female hPSCs X chromosome inactivation was gradually lost with time in culture corresponding to a decrease in XIST expression and an increase in mRNA expression of genes on this chromosome. This type of epigenetic instability is particularly relevant when considering the use of hPSCs in the modelling of X-linked diseases as it could confound results (Mekhoubad et al., 2012).

Further considerations for induced pluripotent stem cells

The issues discussed so far regarding genetic and epigenetic change in culture are similar for both embryonic and induced pluripotent stem cells (Seth et al., 2011). However, there are differences in hiPSCs that present further sources of genetic change in these cells.

hiPSCs differ from hESCs in the fact that they are reprogrammed from somatic tissue. Originally, concerns were raised regarding the use of a retroviral vector for reprogramming (Takahashi et al., 2007) since the integration of the transgene can produce insertion mutations, and insertional mutagenesis has previously been seen to cause serious adverse effects in a gene therapy attempt (Hacein-Bey-Abina et al.,

2003). Attempts to address this issue include the development of reprogramming methods using an episome vector. This is able to replicate extrachromosomally allowing reprogramming without integration. Furthermore, both vector and transgene can then be eliminated via drug selection (Yu et al., 2009).

Mutations that may have been induced during reprogramming have indeed been reported to occur in early passages of iPSC (Hussein et al., 2011), perhaps due to increased replicative stress resulting from forced overexpression of reprogramming factors (Ruiz et al., 2015). However, a comparison of hiPSC lines derived by retroviral or episomal reprogramming showed no significant difference in the frequency of karyotype abnormalities (Taapken et al., 2011). Further, detailed DNA sequence comparisons of parental somatic cells and hiPSC derived from them have indicated that many, if not all of the mutations detected in the hiPSC pre-existed in the parental somatic cells (Gore et al., 2011, Ji et al., 2012, Rouhani et al., 2016) Due to the inefficiency of reprogramming, hiPSC lines usually have a clonal origin and therefore genetic change in just a single parental cell, not detectable in the bulk population because of limited sensitivity of the sequencing methods, could be carried through and mistakenly identified as 'new' genetic variation when the hiPSC culture is compared to the parent culture as a whole (Young et al., 2012). Nevertheless, independent of 'mutations of origin', i.e. those there were present in the parental cells or may have been induced during reprogramming, hiPSC do tend to acquire the same common variants seen in hESC during prolonged culture (Amps et al., 2011, Taapken et al., 2011).

Monitoring genetic change

Monitoring hPSC cultures is important in the laboratory to ensure genetic change does not affect experimental results. It is also vital in clinical applications to ensure cells carrying potentially harmful genetic variation are not introduced into

patients. A number of techniques are available to detect genetic change with some screening the whole genome indiscriminately while others use probes targeted to known loci. The development of single cell based techniques makes it feasible to detect genetic change occurring in only a small minority of cells. All methods, however, have limitations and therefore judgement is required to ensure hPSCs are monitored to an extent that is adequate for their use in either the laboratory or clinical setting.

The traditional, though still highly relevant method for detecting genetic change in cell culture is by assessing the banding pattern of chromosomes in metaphase spreads. This was how some of the earliest genetic changes, such as those on chromosomes 12 and 17 were detected (Figure 1) (Draper et al., 2004). G-banding karyotype analysis has the advantage that it allows assessment of the whole genome for aberrations without any preconceived knowledge, but it is highly labour intensive and analysis usually requires outsourcing to skilled cytogeneticists.

The process for G-banding involves preparing a certain number of metaphase spreads on a slide and scoring a random sample with the assumption that this is representative of the culture as a whole, although it is possible that differential growth patterns or detachment during harvesting of cells in mosaic cultures might distort this assumption. Recently, this assumption and the sensitivity of G-banding was tested systematically using mosaic cultures of hPSC containing known genetic changes at increasing percentages within the population (Baker et al., 2016). The results confirmed that acquired genetic change in hPSC is detected by G-banding at the same frequency as statistically predicted using random sampling. However, sensitivity is limited by cost and practicalities: typically a cytogeneticist might score 30 metaphases, but this will only reliably detect variants that are present in more than 18% of the cells in a mosaic culture (Table 1). A lower limit of around 6% mosaicism requires scoring 100 metaphase spreads, and to reliably detect variants present with less than 1% of the population requires screening over 500 metaphases, a number that is largely impracticable in routine cytogenetic practice. .

G-banding karyotype analysis, even using newly developed automated techniques, is mostly restricted to detecting large genetic aberrations over roughly 5Mb (Steinemann et al., 2013). Therefore, it is rare for small CNVs, such as the common 20q11.21 CNV, to be detected in this way, and typically these require techniques such as SNP array or aCGH based analysis (Amps et al., 2011). The potential of this CNV to be harmful is still unknown. However, as described, its anti-apoptotic property is known to confer a growth advantage and so any plans clinical application involving hPSCs should take account of the inability of karyotype analysis to detect this CNV.

Small CNVs such as that at 20q11.21 can be detected using probe based screening strategies, for example FISH. However, FISH suffers from many of the same issues as G-banding in that it is labour intensive. Also, in practice, it has a limit of detection of around 5% due to false negatives, particularly in the case of tandem duplications when the signals from each copy may overlap and only one copy of the CNV is scored.

To overcome sampling issues it may be possible to combine FISH with flow cytometry in order to conduct a high throughput screen. This interphase chromosome flow-FISH method has been tested on blood samples of myelodysplastic syndrome patients who often present with chromosome 7 monosomy (Keyvanfar et al., 2012). The study found the technique to reliably identify chromosome 7 monosomy without the need for laborious slide analysis. The automated flow cytometry process also allows the screening of thousands of cells at once making it less likely that a genetic aberration will go undetected due to small sample size. Furthermore the technique also provides a quantitative measure of the extent of aneuploidy in the sample.

Recently a quantitative polymerase chain reaction (qPCR) based method has been developed which allows the detection of CNVs based on the comparison of PCR products using primers selected for target and reference regions (Baker et al., 2016). This technique was able to detect CNVs for chromosomes 12, 17 and 20 with a lower detection limit of 10%. This qPCR method provides a very useful technique for routinely

checking laboratory cultures for known common genetic change. However, both qPCR and FISH require pre-existing knowledge of genetic change in order to design primers or probes respectively. This is perhaps not on its own sufficient for clinical applications since we do not yet know the full range of genetic change in hPSC culture or its ability to cause harm, and so a more unbiased screening method may have to be employed as well.

Another powerful genome wide screening method is SNP analysis whereby CNVs are revealed by the increase or decrease in nearby SNP markers detected by microarray platforms. A alternative is comparative genomic hybridisation arrays (aCGH), in which the comparison of sample to reference DNA is more integrated (Rassekh et al., 2008). By hybridizing differentially probed reference and test samples to a microarray the fluorescent ratios of each can be calculated. A ratio of 0 indicates normal or diploid condition whereas ratios of -1 or +1 indicate a loss or gain respectively for that region. This technique has already proved powerful in the field of oncology (Pinkel and Albertson, 2005). Neither SNP array or aCGH approaches require previous knowledge of the genetic change that may be present in a cell population. Furthermore, smaller CNVs of below 5Mb in length can be detected by SNP arrays with the resolution only limited by the distance between SNP markers. The usefulness of this technique has been demonstrated in the screening of 125 hESC lines revealing that greater than 20% carried 20q11.21 CNVs that went largely undetected by karyotype analysis (Amps et al., 2011). However, although SNP based techniques are more precise in terms of the size of CNV they can detect, they do not provide improved sensitivity in detecting CNVs present in only a minority of cells. In mixing tests the ability of SNP microarray analysis to detect chromosome 8 trisomy became unreliable when it was present in less than 10% of the population (Cross et al., 2007). Similar testing of aCGH revealed that the smallest CNVs were only detected in 10-15% cultures (Valli et al., 2011). Another limitation to SNP based analysis is data interpretation. The sensitivity of the method for detecting small genetic changes means that numerous CNVs across all

chromosomes are identified during screening (Amps et al., 2011). However the majority of these will be stochastic in nature and will not produce a significant cellular phenotype. It is therefore a challenge to distinguish the relevant results from the background noise.

Next generation sequencing (NGS) has revolutionised genome research and is increasingly used for the detection of structural variants. Many NGS approaches produce millions of short sequencing reads. By assuming that the distribution of these reads is random over the genome it is possible to infer duplications or deletions from areas that do not follow this trend (Tattini et al., 2015). However, as with other techniques, NGS often fail to detect low level variants of a mosaic population that are hidden by the normal signal. For example, in a study of tumor samples, a coverage as high as 10,000X was required to confirm the presence rare variants (Griffith et al., 2015), Drawing parallels from this highlights the difficulty of sensitively detecting low level mosaicism in cultures of hPSCs. Also, sequencing of repetitive regions is still limited and sequencing or detection of the complete range of genetic variants may often require multiple strategies and sequencing approaches (Tattini et al., 2015).

From these examples it is apparent that there are numerous methods for monitoring genetic change in hPSCs. However, none alone fulfil all of the requirements of a robust detection system. Karyotype analysis is still the best validated and most widely used technology in clinical applications (Whiting et al., 2015) and detects large aberrations, such as those of chromosomes 12 and 17, but we know that significant small CNVs can be missed. Probes for well characterised small CNVs, such as 20q11.21, allow FISH analysis to extend the range of known genetic change that can be detected, but this requires prior knowledge of the CNVs to be assayed. In laboratory applications these techniques may be too labour intensive for routine assessment and so emerging techniques such as qPCR or interphase flow-FISH with a panel of primers or probes could allow one to screen for common genetic changes. In the case of either

clinical or laboratory applications it is important to recognise the limitations of detection methods and judgement is required to achieve satisfactory monitoring of genetic change in using hPSCs

Minimising genetic change

Pertinent to the discussion of acquired genetic change in hPSC's is what measures can be taken to reduce the rate at which genetic variants appear, recognising that their appearance depends upon two unrelated mechanisms, namely mutation, followed by selection. Since much genetic change occurs through prolonged culture it is important to look closely at current methods of passaging and maintaining stem cells in culture. It is also important to discuss novel ways in which the mutation rate can be reduced and whether we can reduce the selection pressure for potentially harmful genetic change also.

Soon after karyotype abnormalities were first linked to prolonged hPSC culture, investigations into the possible effect of different passaging techniques were conducted. For example one group showed that hESC lines could be maintained with a normal karyotype for prolonged periods using a manual passaging technique (Mitalipova et al., 2005). Furthermore, when these same lines were then switched to either enzymatic or non-enzymatic bulk passaging methods, characteristic genetic changes arose. Indeed a correlation between bulk passaging and karyotype abnormalities was documented in a large scale screen (Amps et al., 2011). Certainly the correlation between bulk passaging methods and acquired genetic change may may reflect the different stresses to which cells are exposed by different passaging techniques but it may also be partly due to the greater number of cells that are transferred in bulk methods. For example in a simulation study, the rate at which abnormal cells came to dominate the culture increased exponentially as the size of the overall cell population was increased (Olariu et al., 2010). This is likely a product of a greater number of cells undergoing individual mutational events, which increases the

likelihood of a cell acquiring an advantageous change. This effect could go some way to explaining the higher occurrence of acquired genetic change in hPSC cultures passaged by bulk methods, as the population size will be greater.

The knowledge that population size affects the appearance of genetic variants in culture may provide an opportunity to modify culturing methods. For example, another finding from the simulation studies by Olariu et al. (2010), was that if the same number of cells was cultured in ten smaller sub-cultures the rate at which abnormal cells appeared was lower than in one single large population. The maintenance of hPSCs in small sub-cultures may therefore be an effective way to minimise the effect of genetic change in culture. Furthermore in the laboratory, if one sub-culture did acquire a significant level of genetic change then it could be easily discarded without abandoning the whole experiment. This may also be a useful consideration clinically since many potential regenerative medicine applications will require a significant number of cells. Therefore hPSCs could be expanded through lots of small sub-cultures before combining to produce the final treatment sample although this may not be cost effective or practical for the needs of clinical scale up.

Another consideration regarding hPSC maintenance is how much selection pressure is created by culture methods. It has been documented that during the dissociation of hESC clumps during passaging a large amount of apoptosis occurs (Watanabe et al., 2007) and it was estimated that roughly 90% of cells are lost between each passage (Olariu et al., 2010). This therefore greatly increases the selection pressure for cells carrying genetic change that confers a growth advantage. Increasing the efficiency with which cells are passaged would reduce this selection pressure and therefore could reduce the occurrence of genetic change in culture. One study showed that a ROCK inhibitor could be used to reduce apoptosis during hESC dissociation which significantly increased colony formation after cell transfer and in recent years the use of ROCK inhibitor during hPSC passaging has become common place (Watanabe et al., 2007).

The predominant mechanism of mutation within hPSC culture is poorly understood, but studies have suggested novel ways by which to reduce the incidence of genetic change. For example oxidative stress is widely implicated in DNA damage and hiPSCs have been documented to have high ROS levels following reprogramming (Ji et al., 2014). Furthermore, supplementing hiPSC cultures with antioxidants such as vitamin C reduced ROS and the cells had a reduced number of de novo CNVs (Ji et al., 2014). The use of antioxidants would likely have a similar effect on the mutation rate in hESC culture as well.

Another possible approach may be to use small molecule treatment in order to select against cells with different behaviour conferred by specific genetic variation. For example one group has demonstrated the increased sensitivity of hPSCs carrying trisomy of chromosome 12 to etoposide, cytarabine hydrochloride and gemcitabine hydrochloride (Ben-David et al., 2014), all DNA replication inhibitors already approved as anti-cancer therapies. Since many characterised hPSC genetic abnormalities confer a growth advantage a similar strategy could be employed in culture to select against these cells.

Assessing the effects of genetic change

Despite the possible avenues for reducing genetic change it will be very difficult to culture hPSCs completely free of genetic alterations. Therefore it is very important that we are able to effectively assess genetic variants to distinguish between the problematic and the harmless.

The well documented chromosome 12 and 17 abnormalities can confer a growth advantage and, because of their large scale, they can cause aberrant expression of multiple genes in hPSCs (Draper et al., 2004). Gene expression data from TGCT show copy number increases along chromosome 17q (Skotheim et al., 2002) and isochromosome 12 is used as a clinical marker for TGCT (Bosl et al., 1989). Furthermore investigators have reported a hESC carrying chromosome 12 gains with

neoplastic properties (Werbowetski-Ogilvie et al., 2009). Together these studies suggest that chromosome 12 and 17 abnormalities are unacceptable and therefore all clinical applications of hPSCs should require exclusion of these variants. Most clinical trials include G banding karyotype screens so that, these large chromosomal abnormalities can be excluded with high confidence, providing a satisfactory number of metaphase spreads are analysed. However the question as to when and how often clinically destined samples should be analysed is still unresolved.

Large genetic variation detected at the karyotype level will usually not be acceptable for clinical use. However, a problem arises when considering smaller sub-chromosomal CNVs. The 20q11.21 CNV confers a growth advantage to cells in a similar manner to that associated with chromosome 12 and 17 abnormalities. Therefore one would expect this to be a CNV that needs exclusion during clinical applications. This could be achieved using FISH analysis with a probe specific to the 20q11.21 region. Furthermore, a spectrum of probes could be developed to screen cells for known CNVs. However, genome wide SNP analysis reveals a vast array of CNVs of a similar size to the 20q11.21 (Amps et al., 2011). The problem therefore is that currently we have more data than knowledge. Of the array of CNVs detected using genome wide techniques which, if any, are potentially harmful, either because the variant promotes transformation and the development of cancer, or because it affects the function of the derivative cells to be used for therapy? In either case the answer will be dependent upon the types of derivative cells produced. For example, the potential for converting non-dividing derivative cells such as cardiomyocytes to malignant derivatives is likely to be substantially less than for differentiated cells that still retain proliferative potential, such as hepatocytes.

Clues to the possible consequences for malignancy of genetically variant hPSC derivatives requiremay be obtained from the various cancer genome databases that are now being developed, for example from the International Cancer Genome Consortium (http://icgc.org/). However, direct assessment of malignant potential will

requirein vivo studies. For example, in one study investigators took a hESC line harbouring the 20q11.21 CNV with high proliferative capacity and growth factor independence (Werbowetski-Ogilvie et al., 2009), and transplanted neural derivatives into mice where they formed tumours (Werbowetski-Ogilvie et al., 2012). Similar studies where other recurrent CNVs are tested in vivo could help in the assessment of hPSC genetic variation.

Critical effects of genetic variants on cell function will necessarily have to be tailored to the specific cell types being produced, and could involve either in vivo or in vitro studies as appropriate. For example a vital function of cardiomyocytes is their characteristic calcium handling, which has been used to compare hiPSC derived cardiomyocytes to somatic cells (Hwang et al., 2015). Similar studies with hPSC's carrying a particular CNV could reveal whether the genetic variation disrupts the function of the specialized derivative. This would be extremely important for validating hPSC's as developmental and disease models..

Conclusion

Acquired genomic change is a concern for both its potential to confound the results of basic research and to jeopardise the safety of clinical applications. Despite this, trials using pluripotent stem cell products are in progress. The first such trial was launched by Geron in 2010 which aimed to use oligodendrocyte hESC derivatives to treat spinal cord injury. It was discontinued in 2011 due to financial constraints, but a follow up of the patients occurred at three years (Asterias, 2014). Cardiac progenitors from hESCs have also been used in a trial on heart failure (Menasché et al., 2015). A number of hESC based trials for macular degeneration are also underway including studies launched by Pfizer (Pfizer and London, 2012) and Ocata therapeutics (Schwartz et al., 2015). Thus far, no adverse effects relating to genomic change have been reported in any of these trials. However, it is important to remain vigilant.

Monitoring genetic change has different requirements for specific applications. In basic research efficient, affordable methods are likely to be employed so that they can be applied routinely. Promising techniques, utilizing qPCR and flow cytometry, are therefore likely to be important developments. Monitoring genetic change for clinical applications is likewise changing. For example, in the earliest trial aimed at treating macular degeneration, whilst a normal karyotype was confirmed, further high-resolution techniques were not used (Schwartz et al., 2012). However, in a more recent trial, FISH analysis using probes for loci on chromosomes 12, 17 and 20 was employed to screen for well characterized changes associated with hPSC culture (Menasché et al., 2015). As our knowledge of genomic variation grows, additional probes could be added to this list to exclude other genetic changes. An argument can be made that the technology to indiscriminately screen the whole genome for small CNVs is available in the form of NGS, aCGH and SNP analysis and so should be used. However the difficulty of defining what is a significant genetic change and what is part of normal variation presents a difficulty.

As discussed, one method to assess the significance of genetic variation is through in vivo studies. This already forms a major step in bringing any stem cell based treatment to the clinic. The first macular degeneration trial was preceded by pre-clinical studies in 45 rats which confirmed the safety of the hPSC derived treatment in vivo 12(Li et al., 2012). Since macular degeneration is a disease of the eyes the treatment area is relatively small. This meant the same number of cells ($5x10^4$) could be tested in the model as was used in the human trial (Schwartz et al., 2012). A problem that may arise as hPSC based treatments for larger organs are developed is that the number of cells required will increase. It may, therefore, not be feasible to test the same number of cells in some model organisms due to the relative size of the organ. This is an important consideration because much acquired genetic change occurs during prolonged passage. Therefore, if pre-clinical trials are performed using a smaller

number of cells then the possibility arises that in the human trial the extended culture time required to produce the required cell number will introduce more genetic change.

As hPSCs continue to be used it is likely that protocols will be adjusted to minimize genetic change. For example, a recently launched trial using hiPSC derived retinal pigment epithelium to treat macular degeneration (David, 2014) was put on hold due to the detection of a cancer related mutation in the hiPSC sample (Andy, 2015). This change was not detected in the original skin cells, so could either have been present at undectable levels or were caused by the reprogramming procedure (Andy, 2015). The risk associated with this reprogramming technique is likely to lead to an increased movement towards non-integrative reprogramming techniques such as episomal vectors (Yu et al., 2009). Splitting hPSC cultures into smaller sub-cultures, reducing selection pressure, and using antioxidants may also help to reduce the occurrence of acquired genetic change in culture.

Encouragement can be taken from the lack of adverse effects in human trials using hPSCs to this point. However, it is imperative that this remains the case with future trials for both the safety of patients and to prevent stalling of hPSC applications. This aim will be aided by continual consideration of the monitoring, minimizing and assessment of genomic variation in the context of both basic research and clinical applications.

ALLEGRUCCI, C., WU, Y. Z., THURSTON, A., DENNING, C. N., PRIDDLE, H., MUMMERY, C. L., WARD-VAN OOSTWAARD, D., ANDREWS, P. W., STOJKOVIC, M., SMITH, N., PARKIN, T., JONES, M. E., WARREN, G., YU, L., BRENA, R. M., PLASS, C. & YOUNG, L. E. 2007. Restriction landmark genome scanning identifies culture-induced DNA methylation instability in the human embryonic stem cell epigenome. *Hum Mol Genet,* 16**,** 1253-68.

AMPS, K., ANDREWS, P. W., ANYFANTIS, G., ARMSTRONG, L., AVERY, S., BAHARVAND, H., BAKER, J., BAKER, D., MUNOZ, M. B., BEIL, S., BENVENISTY, N., BEN-YOSEF, D., BIANCOTTI, J. C., BOSMAN, A., BRENA, R. M., BRISON, D., CAISANDER, G., CAMARASA, M. V., CHEN, J., CHIAO, E., CHOI, Y. M., CHOO, A. B., COLLINS, D., COLMAN, A., CROOK, J. M., DALEY, G. Q., DALTON, A., DE SOUSA, P. A., DENNING, C., DOWNIE, J., DVORAK, P., MONTGOMERY, K. D., FEKI, A., FORD, A., FOX, V., FRAGA, A. M., FRUMKIN, T., GE, L., GOKHALE, P. J., GOLAN-LEV, T., GOURABI, H., GROPP, M., LU, G., HAMPL, A., HARRON, K., HEALY, L., HERATH, W., HOLM, F., HOVATTA, O., HYLLNER, J., INAMDAR, M. S., IRWANTO, A. K., ISHII, T., JACONI, M., JIN, Y., KIMBER, S., KISELEV, S., KNOWLES, B. B., KOPPER, O., KUKHARENKO, V., KULIEV, A., LAGARKOVA, M. A., LAIRD, P. W., LAKO, M., LASLETT, A. L., LAVON, N., LEE, D. R., LEE, J. E., LI, C., LIM, L. S., LUDWIG, T. E., MA, Y., MALTBY, E., MATEIZEL, I., MAYSHAR, Y., MILEIKOVSKY, M., MINGER, S. L., MIYAZAKI, T., MOON, S. Y., MOORE, H., MUMMERY, C., NAGY, A., NAKATSUJI, N., NARWANI, K., OH, S. K., OLSON, C., OTONKOSKI, T., PAN, F., PARK, I. H., PELLS, S., PERA, M. F., PEREIRA, L. V., QI, O., RAJ, G. S., REUBINOFF, B., ROBINS, A., ROBSON, P., ROSSANT, J., SALEKDEH, G. H., SCHULZ, T. C., et al. 2011. Screening ethnically diverse human embryonic stem cells identifies a chromosome 20 minimal amplicon conferring growth advantage. *Nat Biotechnol,* 29**,** 1132-44.

ANDREWS, P. W. 2002. From teratocarcinomas to embryonic stem cells. *Philos Trans R Soc Lond B Biol Sci,* 357**,** 405-17.

ANDY, C. 2015. Mutation alert halts stem cell trial to cure blindness. New Scientist: New Scientist.

ASTERIAS, B. 2014. *Asteria Biotherapeutics, Inc. Announces New Results from First-in-Man Clinical Trial of a Cell Therapy Derived from Embryonic Stem Cells* [Online]. http://asteriasbiotherapeutics.com/asterias-biotherapeutics-inc-announces-new-results-from-first-in-man-clinical-trial-of-a-cell-therapy-derived-from-embryonic-stem-cells/: Asteria Biotherapeutics. [Accessed May 22, 2014 2015].

AVERY, S., HIRST, A. J., BAKER, D., LIM, C. Y., ALAGARATNAM, S., SKOTHEIM, R. I., LOTHE, R. A., PERA, M. F., COLMAN, A., ROBSON, P., ANDREWS, P. W. & KNOWLES, B. B. 2013. BCL-XL mediates the strong selective advantage of a 20q11.21 amplification commonly found in human embryonic stem cell cultures. *Stem Cell Reports,* 1**,** 379-86.

BAKER, D., HIRST, A. J., GOKHALE, P. J., JUAREZ, M. A., WILLIAMS, S., WHEELER, M., BEAN, K., ALLISON, T. F., MOORE, H. D., ANDREWS, P. W. & BARBARIC, I. 2016. Detecting Genetic Mosaicism in Cultures of Human Pluripotent Stem Cells. *Stem Cell Reports,* 7**,** 998-1012.

BAKER, D. E., HARRISON, N. J., MALTBY, E., SMITH, K., MOORE, H. D., SHAW, P. J., HEATH, P. R., HOLDEN, H. & ANDREWS, P. W. 2007. Adaptation to culture of human embryonic stem cells and oncogenesis in vivo. *Nat Biotechnol,* 25**,** 207-15.

BLUM, B., BAR-NUR, O., GOLAN-LEV, T. & BENVENISTY, N. 2009. The anti-apoptotic gene survivin contributes to teratoma formation by human embryonic stem cells. *Nat Biotechnol,* 27**,** 281-7.

BOSL, G. J., DMITROVSKY, E., REUTER, V. E., SAMANIEGO, F., RODRIGUEZ, E., GELLER, N. L. & CHAGANTI, R. S. 1989. Isochromosome of the short arm of chromosome 12: clinically useful markers for male germ cell tumors. *J Natl Cancer Inst,* 81**,** 1874-8.

BUTLER, M. G. 2009. Genomic imprinting disorders in humans: a mini-review. *J Assist Reprod Genet,* 26**,** 477-86.

COWAN, C. A., KLIMANSKAYA, I., MCMAHON, J., ATIENZA, J., WITMYER, J., ZUCKER, J. P., WANG, S., MORTON, C. C., MCMAHON, A. P., POWERS, D. & MELTON, D. A. 2004. Derivation of embryonic stem-cell lines from human blastocysts. *N Engl J Med,* 350**,** 1353-6.

CROSS, J., PETERS, G., WU, Z., BROHEDE, J. & HANNAN, G. N. 2007. Resolution of trisomic mosaicism in prenatal diagnosis: estimated performance of a 50K SNP microarray. *Prenat Diagn,* 27**,** 1197-204.

DAVID, C. 2014. Japanese woman is first recipient of next-generation stem cells. Nature News: Nature Publishing Group.

DESMARAIS, J. A., HOFFMANN, M. J., BINGHAM, G., GAGOU, M. E., MEUTH, M. & ANDREWS, P. W. 2012. Human embryonic stem cells fail to activate CHK1 and commit to apoptosis in response to DNA replication stress. *Stem Cells,* 30**,** 1385-93.

DRAPER, J. S., SMITH, K., GOKHALE, P., MOORE, H. D., MALTBY, E., JOHNSON, J., MEISNER, L., ZWAKA, T. P., THOMSON, J. A. & ANDREWS, P. W. 2004. Recurrent gain of chromosomes 17q and 12 in cultured human embryonic stem cells. *Nat Biotechnol,* 22**,** 53-4.

FERGUSON-SMITH, A. C. 2011. Genomic imprinting: the emergence of an epigenetic paradigm. *Nat Rev Genet,* 12**,** 565-75.

FIALKOW, P. J., JACOBSON, R. J. & PAPAYANNOPOULOU, T. 1977. Chronic myelocytic leukemia: clonal origin in a stem cell common to the granulocyte, erythrocyte, platelet and monocyte/macrophage. *Am J Med,* 63**,** 125-30.

GORE, A., LI, Z., FUNG, H. L., YOUNG, J. E., AGARWAL, S., ANTOSIEWICZ-BOURGET, J., CANTO, I., GIORGETTI, A., ISRAEL, M. A., KISKINIS, E., LEE, J. H., LOH, Y. H., MANOS, P. D., MONTSERRAT, N., PANOPOULOS, A. D., RUIZ, S., WILBERT, M. L., YU, J., KIRKNESS, E. F., IZPISUA BELMONTE, J. C., ROSSI, D. J., THOMSON, J. A., EGGAN, K., DALEY, G. Q., GOLDSTEIN, L. S. & ZHANG, K. 2011. Somatic coding mutations in human induced pluripotent stem cells. *Nature,* 471**,** 63-7.

GRIFFITH, M., MILLER, C. A., GRIFFITH, O. L., KRYSIAK, K., SKIDMORE, Z. L., RAMU, A., WALKER, J. R., DANG, H. X., TRANI, L., LARSON, D. E., DEMETER,

R. T., WENDL, M. C., MCMICHAEL, J. F., AUSTIN, R. E., MAGRINI, V., MCGRATH, S. D., LY, A., KULKARNI, S., CORDES, M. G., FRONICK, C. C., FULTON, R. S., MAHER, C. A., DING, L., KLCO, J. M., MARDIS, E. R., LEY, T. J. & WILSON, R. K. 2015. Optimizing cancer genome sequencing and analysis. *Cell Syst,* 1**,** 210-223.

HACEIN-BEY-ABINA, S., VON KALLE, C., SCHMIDT, M., MCCORMACK, M. P., WULFFRAAT, N., LEBOULCH, P., LIM, A., OSBORNE, C. S., PAWLIUK, R., MORILLON, E., SORENSEN, R., FORSTER, A., FRASER, P., COHEN, J. I., DE SAINT BASILE, G., ALEXANDER, I., WINTERGERST, U., FREBOURG, T., AURIAS, A., STOPPA-LYONNET, D., ROMANA, S., RADFORD-WEISS, I., GROSS, F., VALENSI, F., DELABESSE, E., MACINTYRE, E., SIGAUX, F., SOULIER, J., LEIVA, L. E., WISSLER, M., PRINZ, C., RABBITTS, T. H., LE DEIST, F., FISCHER, A. & CAVAZZANA-CALVO, M. 2003. LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1. *Science,* 302**,** 415-9.

HUMPHERYS, D., EGGAN, K., AKUTSU, H., HOCHEDLINGER, K., RIDEOUT, W. M., BINISZKIEWICZ, D., YANAGIMACHI, R. & JAENISCH, R. 2001. Epigenetic instability in ES cells and cloned mice. *Science,* 293**,** 95-7.

HUSSEIN, S. M., BATADA, N. N., VUORISTO, S., CHING, R. W., AUTIO, R., NÄRVÄ, E., NG, S., SOUROUR, M., HÄMÄLÄINEN, R., OLSSON, C., LUNDIN, K., MIKKOLA, M., TROKOVIC, R., PEITZ, M., BRÜSTLE, O., BAZETT-JONES, D. P., ALITALO, K., LAHESMAA, R., NAGY, A. & OTONKOSKI, T. 2011. Copy number variation and selection during reprogramming to pluripotency. *Nature,* 471**,** 58-62.

HWANG, H. S., KRYSHTAL, D. O., FEASTER, T. K., SÁNCHEZ-FREIRE, V., ZHANG, J., KAMP, T. J., HONG, C. C., WU, J. C. & KNOLLMANN, B. C. 2015. Comparable calcium handling of human iPSC-derived cardiomyocytes generated by multiple laboratories. *J Mol Cell Cardiol,* 85**,** 79-88.

ITZHAKI, I., MAIZELS, L., HUBER, I., ZWI-DANTSIS, L., CASPI, O., WINTERSTERN, A., FELDMAN, O., GEPSTEIN, A., ARBEL, G., HAMMERMAN, H., BOULOS, M. & GEPSTEIN, L. 2011. Modelling the long QT syndrome with induced pluripotent stem cells. *Nature,* 471**,** 225-9.

JI, J., NG, S. H., SHARMA, V., NECULAI, D., HUSSEIN, S., SAM, M., TRINH, Q., CHURCH, G. M., MCPHERSON, J. D., NAGY, A. & BATADA, N. N. 2012. Elevated coding mutation rate during the reprogramming of human somatic cells into induced pluripotent stem cells. *Stem Cells,* 30**,** 435-40.

JI, J., SHARMA, V., QI, S., GUARCH, M. E., ZHAO, P., LUO, Z., FAN, W., WANG, Y., MBABAALI, F., NECULAI, D., ESTEBAN, M. A., MCPHERSON, J. D. & BATADA, N. N. 2014. Antioxidant supplementation reduces genomic aberrations in human induced pluripotent stem cells. *Stem Cell Reports,* 2**,** 44-51.

KEYVANFAR, K., WEED, J., SWAMY, P., KAJIGAYA, S., CALADO, R. T. & YOUNG, N. S. 2012. Interphase Chromosome Flow-FISH. *Blood,* 120**,** e54-9.

KRIKS, S., SHIM, J. W., PIAO, J., GANAT, Y. M., WAKEMAN, D. R., XIE, Z., CARRILLO-REID, L., AUYEUNG, G., ANTONACCI, C., BUCH, A., YANG, L., BEAL, M. F., SURMEIER, D. J., KORDOWER, J. H., TABAR, V. & STUDER, L. 2011. Dopamine neurons derived from human ES cells efficiently engraft in animal models of Parkinson's disease. *Nature,* 480**,** 547-51.

LAURENT, L. C., ULITSKY, I., SLAVIN, I., TRAN, H., SCHORK, A., MOREY, R., LYNCH, C., HARNESS, J. V., LEE, S., BARRERO, M. J., KU, S., MARTYNOVA, M., SEMECHKIN, R., GALAT, V., GOTTESFELD, J., IZPISUA BELMONTE, J. C., MURRY, C., KEIRSTEAD, H. S., PARK, H. S., SCHMIDT, U., LASLETT, A. L., MULLER, F. J., NIEVERGELT, C. M., SHAMIR, R. & LORING, J. F. 2011. Dynamic changes in the copy number of pluripotency and cell proliferation genes in human ESCs and iPSCs during reprogramming and time in culture. *Cell Stem Cell,* 8**,** 106-18.

LI, Y., TSAI, Y. T., HSU, C. W., EROL, D., YANG, J., WU, W. H., DAVIS, R. J., EGLI, D. & TSANG, S. H. 2012. Long-term safety and efficacy of human-induced pluripotent stem cell (iPS) grafts in a preclinical model of retinitis pigmentosa. *Mol Med,* 18**,** 1312-9.

MAYSHAR, Y., BEN-DAVID, U., LAVON, N., BIANCOTTI, J. C., YAKIR, B., CLARK, A. T., PLATH, K., LOWRY, W. E. & BENVENISTY, N. 2010. Identification and classification of chromosomal aberrations in human induced pluripotent stem cells. *Cell Stem Cell,* 7**,** 521-31.

MEKHOUBAD, S., BOCK, C., DE BOER, A. S., KISKINIS, E., MEISSNER, A. & EGGAN, K. 2012. Erosion of dosage compensation impacts human iPSC disease modeling. *Cell Stem Cell,* 10**,** 595-609.

MENASCHÉ, P., VANNEAUX, V., HAGÈGE, A., BEL, A., CHOLLEY, B., CACCIAPUOTI, I., PAROUCHEV, A., BENHAMOUDA, N., TACHDJIAN, G., TOSCA, L., TROUVIN, J. H., FABREGUETTES, J. R., BELLAMY, V., GUILLEMAIN, R., SUBERBIELLE BOISSEL, C., TARTOUR, E., DESNOS, M. & LARGHERO, J. 2015. Human embryonic stem cell-derived cardiac progenitors for severe heart failure treatment: first clinical case report. *Eur Heart J,* 36**,** 2011-7.

MITALIPOVA, M. M., RAO, R. R., HOYER, D. M., JOHNSON, J. A., MEISNER, L. F., JONES, K. L., DALTON, S. & STICE, S. L. 2005. Preserving the genetic integrity of human embryonic stem cells. *Nat Biotechnol,* 23**,** 19-20.

MITSUI, K., TOKUZAWA, Y., ITOH, H., SEGAWA, K., MURAKAMI, M., TAKAHASHI, K., MARUYAMA, M., MAEDA, M. & YAMANAKA, S. 2003. The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. *Cell,* 113**,** 631-42.

NAZOR, K. L., ALTUN, G., LYNCH, C., TRAN, H., HARNESS, J. V., SLAVIN, I., GARITAONANDIA, I., MÜLLER, F. J., WANG, Y. C., BOSCOLO, F. S., FAKUNLE, E., DUMEVSKA, B., LEE, S., PARK, H. S., OLEE, T., D'LIMA, D. D., SEMECHKIN, R., PARAST, M. M., GALAT, V., LASLETT, A. L., SCHMIDT, U., KEIRSTEAD, H. S., LORING, J. F. & LAURENT, L. C. 2012. Recurrent variations in DNA methylation in human pluripotent stem cells and their differentiated derivatives. *Cell Stem Cell,* 10**,** 620-34.

OLARIU, V., HARRISON, N. J., COCA, D., GOKHALE, P. J., BAKER, D., BILLINGS, S., KADIRKAMANATHAN, V. & ANDREWS, P. W. 2010. Modeling the evolution of culture-adapted human embryonic stem cells. *Stem Cell Res,* 4**,** 50-6.

PERRIER, A. L., TABAR, V., BARBERI, T., RUBIO, M. E., BRUSES, J., TOPF, N., HARRISON, N. L. & STUDER, L. 2004. Derivation of midbrain dopamine neurons from human embryonic stem cells. *Proc Natl Acad Sci U S A,* 101**,** 12543-8.

PETERSON, S. E., WESTRA, J. W., REHEN, S. K., YOUNG, H., BUSHMAN, D. M., PACZKOWSKI, C. M., YUNG, Y. C., LYNCH, C. L., TRAN, H. T., NICKEY, K. S.,

WANG, Y. C., LAURENT, L. C., LORING, J. F., CARPENTER, M. K. & CHUN, J. 2011. Normal human pluripotent stem cell lines exhibit pervasive mosaic aneuploidy. *PLoS One,* 6**,** e23018.

PFIZER & LONDON, U. C. 2012. https://clinicaltrials.gov/ct2/show/NCT01691261: Clinical Trials. [Accessed September 22, 2012 2012].

PINKEL, D. & ALBERTSON, D. G. 2005. Array comparative genomic hybridization and its applications in cancer. *Nat Genet,* 37 Suppl**,** S11-7.

RASSEKH, S. R., CHAN, S., HARVARD, C., DIX, D., QIAO, Y. & RAJCAN-SEPAROVIC, E. 2008. Screening for submicroscopic chromosomal rearrangements in Wilms tumor using whole-genome microarrays. *Cancer Genet Cytogenet,* 182**,** 84-94.

REHEN, S. K., YUNG, Y. C., MCCREIGHT, M. P., KAUSHAL, D., YANG, A. H., ALMEIDA, B. S., KINGSBURY, M. A., CABRAL, K. M., MCCONNELL, M. J., ANLIKER, B., FONTANOZ, M. & CHUN, J. 2005. Constitutional aneuploidy in the normal human brain. *J Neurosci,* 25**,** 2176-80.

ROUHANI, F. J., NIK-ZAINAL, S., WUSTER, A., LI, Y., CONTE, N., KOIKE-YUSA, H., KUMASAKA, N., VALLIER, L., YUSA, K. & BRADLEY, A. 2016. Mutational History of a Human Cell Lineage from Somatic to Induced Pluripotent Stem Cells. *PLoS Genet,* 12**,** e1005932.

RUGG-GUNN, P. J., FERGUSON-SMITH, A. C. & PEDERSEN, R. A. 2005. Epigenetic status of human embryonic stem cells. *Nat Genet,* 37**,** 585-7.

RUGG-GUNN, P. J., FERGUSON-SMITH, A. C. & PEDERSEN, R. A. 2007. Status of genomic imprinting in human embryonic stem cells as revealed by a large cohort of independently derived and maintained lines. *Hum Mol Genet,* 16 Spec No. 2**,** R243-51.

RUIZ, S., LOPEZ-CONTRERAS, A. J., GABUT, M., MARION, R. M., GUTIERREZ-MARTINEZ, P., BUA, S., RAMIREZ, O., OLALDE, I., RODRIGO-PEREZ, S., LI, H., MARQUES-BONET, T., SERRANO, M., BLASCO, M. A., BATADA, N. N. & FERNANDEZ-CAPETILLO, O. 2015. Limiting replication stress during somatic cell reprogramming reduces genomic instability in induced pluripotent stem cells. *Nat Commun,* 6**,** 8036.

SCHWARTZ, S. D., HUBSCHMAN, J. P., HEILWELL, G., FRANCO-CARDENAS, V., PAN, C. K., OSTRICK, R. M., MICKUNAS, E., GAY, R., KLIMANSKAYA, I. & LANZA, R. 2012. Embryonic stem cell trials for macular degeneration: a preliminary report. *Lancet,* 379**,** 713-20.

SCHWARTZ, S. D., REGILLO, C. D., LAM, B. L., ELIOTT, D., ROSENFELD, P. J., GREGORI, N. Z., HUBSCHMAN, J. P., DAVIS, J. L., HEILWELL, G., SPIRN, M., MAGUIRE, J., GAY, R., BATEMAN, J., OSTRICK, R. M., MORRIS, D., VINCENT, M., ANGLADE, E., DEL PRIORE, L. V. & LANZA, R. 2015. Human embryonic stem cell-derived retinal pigment epithelium in patients with age-related macular degeneration and Stargardt's macular dystrophy: follow-up of two open-label phase 1/2 studies. *Lancet,* 385**,** 509-16.

SKOTHEIM, R. I., MONNI, O., MOUSSES, S., FOSSÅ, S. D., KALLIONIEMI, O. P., LOTHE, R. A. & KALLIONIEMI, A. 2002. New insights into testicular germ cell tumorigenesis from gene expression profiling. *Cancer Res,* 62**,** 2359-64.

SOLDNER, F., HOCKEMEYER, D., BEARD, C., GAO, Q., BELL, G. W., COOK, E. G., HARGUS, G., BLAK, A., COOPER, O., MITALIPOVA, M., ISACSON, O. &

JAENISCH, R. 2009. Parkinson's disease patient-derived induced pluripotent stem cells free of viral reprogramming factors. *Cell,* 136**,** 964-77.
STEINEMANN, D., GÖHRING, G. & SCHLEGELBERGER, B. 2013. Genetic instability of modified stem cells - a first step towards malignant transformation? *Am J Stem Cells,* 2**,** 39-51.
TAAPKEN, S. M., NISLER, B. S., NEWTON, M. A., SAMPSELL-BARRON, T. L., LEONHARD, K. A., MCINTIRE, E. M. & MONTGOMERY, K. D. 2011. Karotypic abnormalities in human induced pluripotent stem cells and embryonic stem cells. *Nat Biotechnol,* 29**,** 313-4.
TAKAHASHI, K., TANABE, K., OHNUKI, M., NARITA, M., ICHISAKA, T., TOMODA, K. & YAMANAKA, S. 2007. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell,* 131**,** 861-72.
TATTINI, L., D'AURIZIO, R. & MAGI, A. 2015. Detection of Genomic Structural Variants from Next-Generation Sequencing Data. *Front Bioeng Biotechnol,* 3**,** 92.
THOMSON, J. A., ITSKOVITZ-ELDOR, J., SHAPIRO, S. S., WAKNITZ, M. A., SWIERGIEL, J. J., MARSHALL, V. S. & JONES, J. M. 1998. Embryonic stem cell lines derived from human blastocysts. *Science,* 282**,** 1145-7.
VALLI, R., MARLETTA, C., PRESSATO, B., MONTALBANO, G., LO CURTO, F., PASQUALI, F. & MASERATI, E. 2011. Comparative genomic hybridization on microarray (a-CGH) in constitutional and acquired mosaicism may detect as low as 8% abnormal cells. *Mol Cytogenet,* 4**,** 13.
WATANABE, K., UENO, M., KAMIYA, D., NISHIYAMA, A., MATSUMURA, M., WATAYA, T., TAKAHASHI, J. B., NISHIKAWA, S., MUGURUMA, K. & SASAI, Y. 2007. A ROCK inhibitor permits survival of dissociated human embryonic stem cells. *Nat Biotechnol,* 25**,** 681-6.
WERBOWETSKI-OGILVIE, T. E., BOSSÉ, M., STEWART, M., SCHNERCH, A., RAMOS-MEJIA, V., ROULEAU, A., WYNDER, T., SMITH, M. J., DINGWALL, S., CARTER, T., WILLIAMS, C., HARRIS, C., DOLLING, J., WYNDER, C., BOREHAM, D. & BHATIA, M. 2009. Characterization of human embryonic stem cells with features of neoplastic progression. *Nat Biotechnol,* 27**,** 91-7.
WERBOWETSKI-OGILVIE, T. E., MORRISON, L. C., FIEBIG-COMYN, A. & BHATIA, M. 2012. In vivo generation of neural tumors from neoplastic pluripotent stem cells models early human pediatric brain tumor formation. *Stem Cells,* 30**,** 392-404.
WHITING, P., KERBY, J., COFFEY, P., DA CRUZ, L. & MCKERNAN, R. 2015. Progressing a human embryonic stem-cell-based regenerative medicine therapy towards the clinic. *Philos Trans R Soc Lond B Biol Sci,* 370**,** 20140375.
YOUNG, M. A., LARSON, D. E., SUN, C. W., GEORGE, D. R., DING, L., MILLER, C. A., LIN, L., PAWLIK, K. M., CHEN, K., FAN, X., SCHMIDT, H., KALICKI-VEIZER, J., COOK, L. L., SWIFT, G. W., DEMETER, R. T., WENDL, M. C., SANDS, M. S., MARDIS, E. R., WILSON, R. K., TOWNES, T. M. & LEY, T. J. 2012. Background mutations in parental cells account for most of the genetic heterogeneity of induced pluripotent stem cells. *Cell Stem Cell,* 10**,** 570-82.

YU, J., HU, K., SMUGA-OTTO, K., TIAN, S., STEWART, R., SLUKVIN, I. I. & THOMSON, J. A. 2009. Human induced pluripotent stem cells free of vector and transgene sequences. *Science,* 324**,** 797-801.

YU, J., VODYANIK, M. A., SMUGA-OTTO, K., ANTOSIEWICZ-BOURGET, J., FRANE, J. L., TIAN, S., NIE, J., JONSDOTTIR, G. A., RUOTTI, V., STEWART, R., SLUKVIN, I. I. & THOMSON, J. A. 2007. Induced pluripotent stem cell lines derived from human somatic cells. *Science,* 318**,** 1917-20.

Table 1. The sensitivity of detecting karyotypically variant cells in mosaic cultures by G-Banding karyology

Number of Metaphases Scored	Percent Variant Cells Detected with 95% Confidence
20	28%
30	18%
50	13%
60	10%
100	6%
500	< 1%

The table shows, based on statistical sampling theory, the minimum proportion of variant cells that would be detected in mosaic cultures for different numbers of metaphases scored (Baker et al 2016). By screening test cultures with different proportions of variant hESC, the actual sensitivity of G-banding karyology carried out using standard procedures closely matched the expected sensitivity predicted by statistical sampling theory.